Overview: CCS-6 Statistical Sciences Group

Joanne Wendelberger

Christine Anderson-Cook

Dave Higdon

LA-UR-06-4482

Overview: CCS-6, Statistical Sciences Group

- Introduction to CCS-6
  - History/Mission/Vision
  - People
  - Technical Capabilities
  - Customers/Projects
- Technical Mini-Briefs
  - Reliability of a Complex System, Christine Anderson-Cook
  - Combining Computer Models and Experiments, Dave Higdon
- Wrap-Up, Dave Higdon

Statistical Sciences Group

History:

Founded as the Statistics Group in 1967, the group will celebrate its 40th Anniversary in 2007.

Vision:

Achieve excellence in development of techniques for collecting, analyzing, combining, and making inferences from diverse qualitative and quantitative information sets such as experiments, observational studies, computer simulations, and expert judgment.

Statistical Sciences Group

Mission:
Bring statistical reasoning and rigor to multi-disciplinary scientific investigations through development, application, and communication of cutting-edge statistical sciences research.

Action:
Work in partnerships with scientists, engineers and policy makers within and outside the Laboratory to solve problems of national importance.

CCS-6, Statistical Sciences

David Higdon
Group Leader

Joanne Wendelberger
Deputy Group Leader

Yvonne M. Armijo
Administrative Operations Specialist

Vacant
Administrative Specialist

Kenneth Cox
Administrator

Post Docs
Wai F. (Calvin) Chiu
Chris Orum
Margaret Short

GRAs
Alina Kline
Ivan Ramler
Kari Sentz
Brian Weaver

Staff

Christine Anderson-Cook
Thomas L. Burr
Cheryll Faust
Michael L. Fugate
James R. Gattiker
Todd L. Graves
Michael S. Hamada
Geralyn M. Hemphill
Elizabeth J. Kelly
Richard Klamann
Hazel Kutac
Earl Lawrence
Harry F. Martz

Staff (con't.)

Michael McKay
Sarah E. Michalak
Leslie M. Moore
Kary Myers
Richard R. Picard
William H. Press
Christopher S. Reese
Vivian L. Romero
Lawrence O. Ticknor
Diane Tompkins
Scott Vander Wiel
Brian J. Williams
Alyson G. Wilson

System Ethnography and Qualitative Modeling team (SEQM)

Andrew C. Koehler
Benjamin H. Sims
Gregory D. Wilson

Visiting Faculty

Peter J. Bickel
Derek R. Bingham
Arthur Dempster
George T. Duncan
Aparna V. Huurbazar
Carl G. Herndl
Hariharan K. Iyer
Daniel Jeske
Valen E. Johnson
Sallie Keller-McNulty
John C. Kern II
Todd R. La Porte
Chuanhai Liu
Jason Loeppky
J. Stephen Marron
William Q. Meeker
Max D. Morris

Visiting Faculty (con't.)

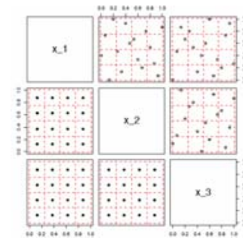
Timothy J. Robinson
Thomas J. Santner
David W. Scott
Nozer Singpuralla
Randy R. Sitter
Paul L. Speckman
Sara L. Stokes
Steven K. Thompson
Stephen Vardeman
Huaiqing Wu

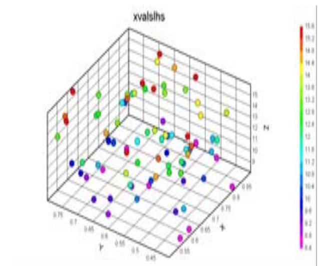
Consultant/Guest Scientist

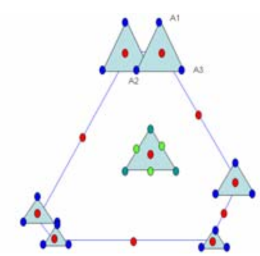
Arthur Koehler
Jerome Morzinski
Robert D. Ryne

Technical Capabilities

- Data Analysis and Computational Statistics
- Theory and Methods for Computer Model Evaluation
- Monte Carlo Methods
- Reliability
- System Ethnography and Qualitative Modeling
- Information Integration Technology
- Uncertainty Quantification, Statistical Bounding
- Design and Analysis of Experiments
- Biological Sciences Applications







CCS-6 Projects and Customers

- Diverse project/customer base.
- Provide statistical expertise in a variety of customer relationships.
- Many collaborative projects in both lead and support roles.

Selected Projects Led by CCS-6

- Design Agency System Point of Contact for Reliability (DASPOC) work for Nuclear Weapons Program
- Enhanced Reliability Modeling for Nuclear Weapons Program
- Computer Model Evaluation for Weapons Physics for X Division
- Joint Munitions Project – DoD
- Sampling Strategies for BIONET for DTRA/DHS
- Model Evaluation Methods for Procter & Gamble
- Shelf-life Modeling for Formulated Products for Procter & Gamble
- LDRD – Design of Experiment Construction and Assessment
- Institutional Program Development – Design and Analysis of Experiments and Sampling
- TOW Missile Analyses for Marine Corps Programs
- Ballistic Missile Defense for Missile Defense Agency

Selected Projects Supported by CCS-6

- Reliability, Statistical Support for ESA Surveillance Team
- Significant Finding Investigations for the Nuclear Weapons Program
- Statistical sampling, design, and analysis for W76 Life Extension Program
- Statistical Studies for CSA MTE
- Biosense for CDC
- Plutonium Metal Exchange Program
- Amplified Fragment Length Polymorphisms Studies with B Division
- Non-Proliferation detection with C-INC
- Remeasurement Database and Propagation of Variance Modeling for Nuclear Safeguards with NMT
- LDRD – Metabalomic Studies collaboration with B Division
- Biological Risk Assessment Team for Homeland Security led by D-4
- Infrastructure Issues for NISAC, CIP/DSS
- Statistical analyses for Pit Production for NMT

Stockpile Reliability Assessment - Summary

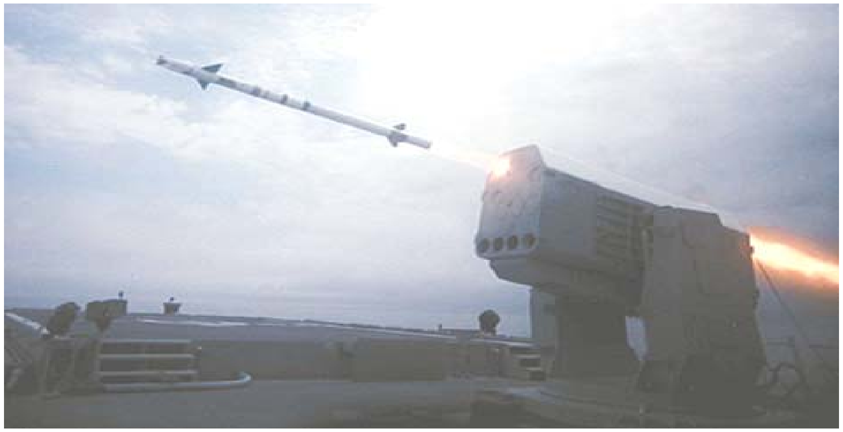
- Goal/Objective: The goal of this project is to develop a dependable and cost-effective suite of statistical methodologies and tools to assess the reliability of weapons stockpiles.

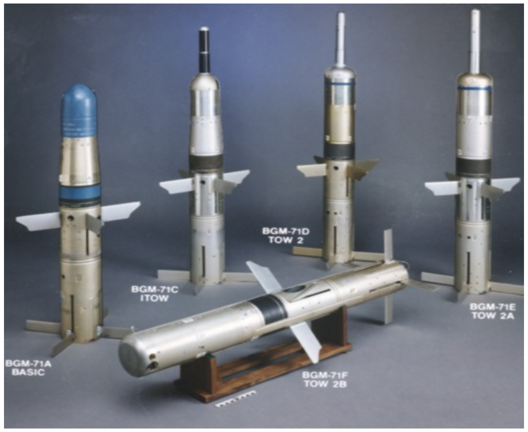
- Approach
  - Methodological development
    - information integration
    - uncertainty quantification with heterogeneous data
  - Applications collaboration
  - Tool development
    - software for rapid development of systems and statistical models

Collaborators and Customers

- DoD
  - MCPD Fallbrook (TOW)
  - NSWC Corona (RAM, ESSM)
  - NSWC Yorktown (AMRAAM)
  - AMCOM/RDEC (Stinger)

- DOE
  - LANL Enhanced Surveillance Campaign
  - LANL Core Surveillance





Group TSMs Involved with Work

- Christine Anderson-Cook
- Cheryll Faust
- Todd Graves
- Michael Hamada
- Richard Klamann
- Andrew Koehler
- Earl Lawrence
- Harry Martz
- Shane Reese
- Benjamin Sims
- Scott Vander Wiel
- Alyson Wilson
- Greg Wilson

The fundamental question is how to assess stockpiles as they change over time.

- Stockpiles change over time due to materials degradation, life-extension programs, maintenance, use, and other factors.

- Assessment requires
  - the development of system models that capture parts, functions, dynamics, and interactions
  - the integration of multiple data sources, including historical data, surveillance testing, accelerated life testing, computer model output, and materials characterization.

The (growing) Challenge

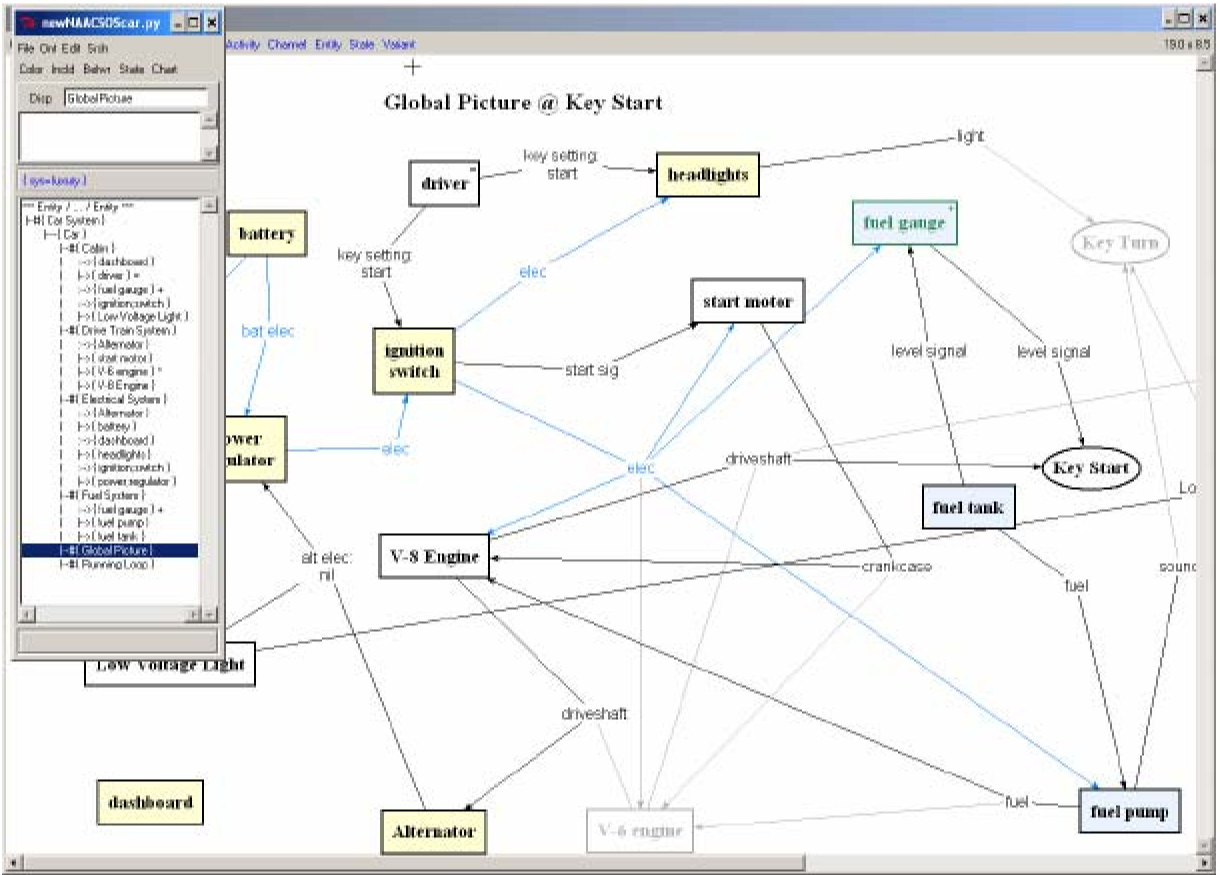
Suppose that we are trying to assess a stockpile that has

- Multiple variants,
- Multiple data sources,
- Distributed expertise,
- *Limits on functional testing*

and that we want

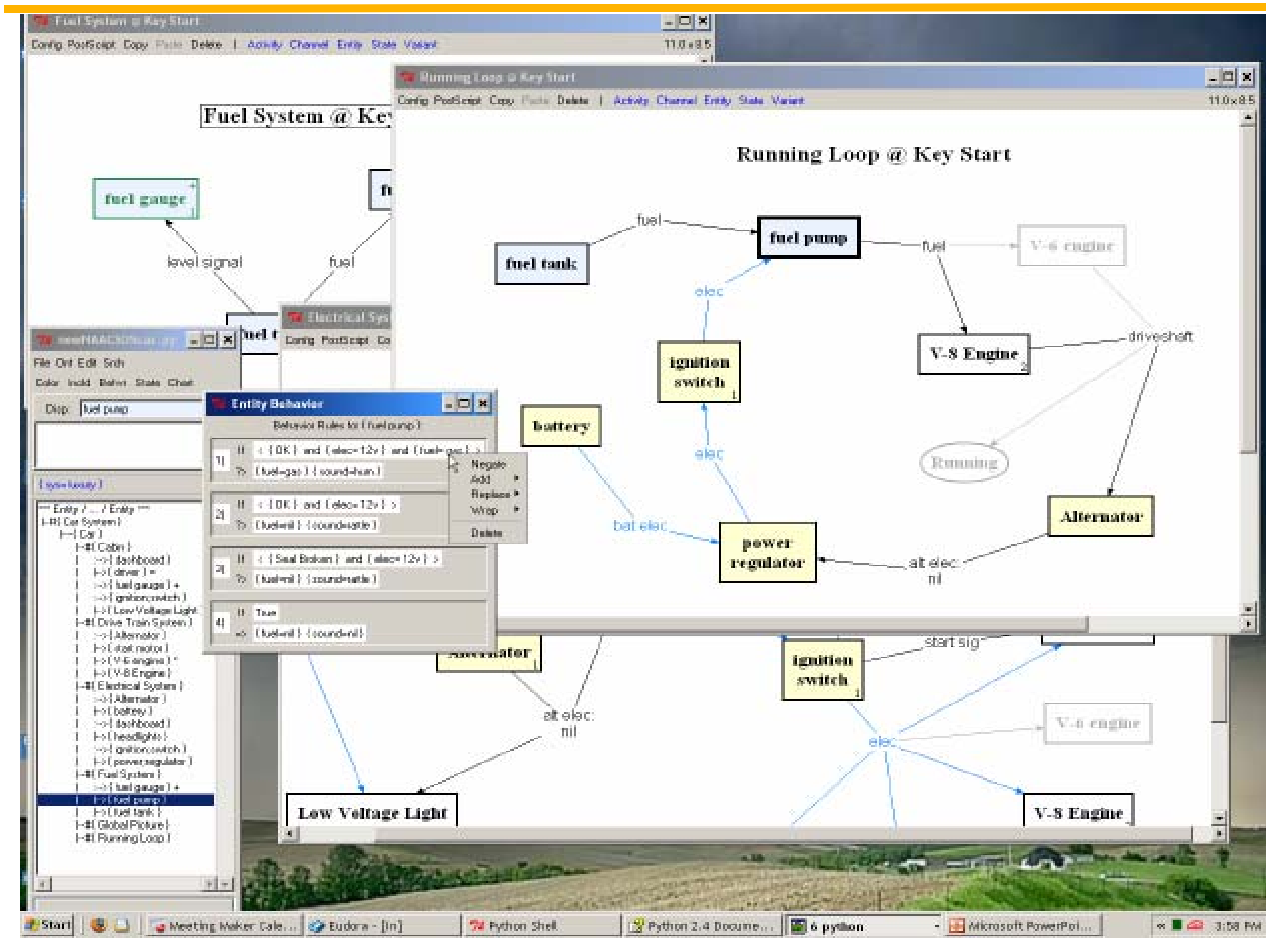
- A numerical estimate of current reliability and performance based on individual and group characteristics,
- A prediction of how reliability and performance change over time,
- Uncertainties on the estimates and predictions,
- A system description that captures stockpile environments and use dynamics.

Combine all available information to understand uncertainties in system reliability and performance.



- Data is often available from many different experiments: flight tests, component tests, accelerated life tests.
- GROMIT allows us to understand what the data tells us about the system.
- We also develop statistical methods to formally combine the information into a unified system reliability estimate.

GROMIT allows us to combine information from different experts into an integrated system view.



- Different subject matter experts understand different parts of the system.
- GROMIT highlights potential differences in system assumptions and understanding from various experts, to create a more accurate system representation.
- Effective assessment requires an integrated system view.

Overview of Statistical Model

- Model contains a reliability distribution for each component, as well as how the components are combined to give system
- For basic model:
  - Reliability distribution for each component as a function of age will be estimated from the data and any expert knowledge that we wish to incorporate
  - Components combine into whole system (serially or with redundancy to reflect design of system)
- Other aspects:
  - Component reliability will be estimated by using both flight and component quality assurance measures
  - Different variants of systems are possible

Bayesian Analysis

- Capability to easily incorporate expert knowledge about reliability for individual components, through informative priors
- Using special-purpose MCMC programming packages, YADAS
  - http://yadas.lanl.gov
  - Control over algorithm choices
  - “Solves” broader class of models
- Analysis could also be programmed in other languages as well (eg. R, S-Plus, WinBugs)
- Computationally quite intensive

Translation of Data – Full-System Data

- Need to translate flight successes and failures into information about the individual components of the system

SN	Age (months)	Result	Failure Mode
U00866	110	Failure	Missile Battery Fails in Flight
U00867	56	Success	
U00868	87	Success	
...	...	...	
U00843	33	Success	
U00858	91	Success	
U00818	103	Failure	Degraded Roll
U00814	41	Failure	Hangfire
U00803	74	Failure	Unguided Flight

Activity	Failure Mode	Related Hardware	Possible Root Cause
RAM Designation [From Missile Present to ITL (p A-22)]	Missile Not Detected	Ship	Error in Ship Controls
		Ship, Launcher	Error in Ship to Launcher Interface
		Launcher	Error in Launcher
		Umbilical	Error in Launcher to GMRP Interface
RAM Launch [From Successful Missile Ready to Umbilical Separation]	Misfire	Rocket Motor	No/Low RM Thrust
	Hangfire	Canister: Hold Back Latch	HBL not retracted
	Dud	Rocket Motor	RM not fired
		Canister: Squibs	Failed Squibs
		Ship to Launcher	Miscommunition, No Signal
		Launcher to Missile (Umbilical)	Open Wires, Failed Umbilical
	Launch Cover Eject Fails	Canister Squibs	Squibs Failed, Covers did not completely separate
	Low RM Thrust	Rocket Motor	Aged Propellant, Propellant not ignited
	Degraded Roll	Rocket Motor	Low RM Thrust

Activity	Failure Mode	Related Hardware	Possible Root Cause
RAM Designation [From Missile Present to ITL (p A-22)]	Missile Not Detected	Ship	Error in Ship Controls
		Ship, Launcher	Error in Ship to Launcher Interface
		Launcher	Error in Launcher
		Umbilical	Error in Launcher to GMRP Interface

RAM Launch [From Successful Missile Ready to Umbilical Separation]	Misfire	Rocket Motor	No/Low RM Thrust
	Hangfire	Canister: Hold Back Latch	HBL not retracted
	Dud	Rocket Motor	RM not fired
		Canister: Squibs	Failed Squibs
		Ship to Launcher	Miscommunition, No Signal
		Launcher to Missile (Umbilical)	Open Wires, Failed Umbilical
	Launch Cover Eject Fails	Canister Squibs	Squibs Failed, Covers did not completely separate
	Low RM Thrust	Rocket Motor	Aged Propellant, Propellant not ignited
	Degraded Roll	Rocket Motor	Low RM Thrust

Legend
S = Success
F = Failure
PF = Possible Failure
? = Status Not Tested

At least one of these failed

Once a system fails, no info about components in later phases

Specific component failed

	P1:Missile Assign			P2: Wake Up				P3: Launch			P8: Direct Strike			
Failure Modes	C1	C2	C3	C4	C5	C6	C7	C8	C9	…	C27	C28	C29	C30
Missile Not Detected	PF	PF	PF	?	?	?	?	?	?		?	?	?	?
Misfire	S	S	S	S	S	S	S	F	S		?	?	?	?
Hangfire	S	S	S	S	S	S	S	S	F		?	?	?	?
Dud	S	S	S	S	S	S	S	PF	PF		?	?	?	?
Degraded Pitch	S	S	S	S	S	S	S	S	S		?	?	?	?
…														
Success	S	S	S	S	S	S	S	S	S		S	S	S	S

Translation of Data – Component Quality Assurance Data

C1			
Age	Value	Lower Spec	Upper Spec
1	12.6	10	15
1	13.1	10	15
2	13.2	10	15
2	14.1	10	15
2	13.4	10	15
3	13.7	10	15
…	…	…	…
6	14.8	10	15

C2			
Age	Value	Lower Spec	Upper Spec
1	1.25	1	2
1	1.33	1	2
2	1.51	1	2
2	1.26	1	2
2	1.44	1	2
3	1.24	1	2
…		…	…
6	1.37	1	2

• • •

C4		
Age	Value	Upper Spec
1	3.12	4
1	3.17	4
2	3.22	4
2	3.41	4
2	3.15	4
3	3.28	4
…	…	…
6	3.18	4

C1		
Age	Value	Lower Spec
1	12.3	10
1	14.3	10
2	15.1	10
2	11.1	10
2	9.7	10
3	10.3	10
…	…	…
6	14.2	10

- Some components may not have any quality assurance data
- Some components may have multiple measures
- Specification limits can be Upper and Lower, Lower Only, Upper Only

Integrating Components of Model into Unified Analysis

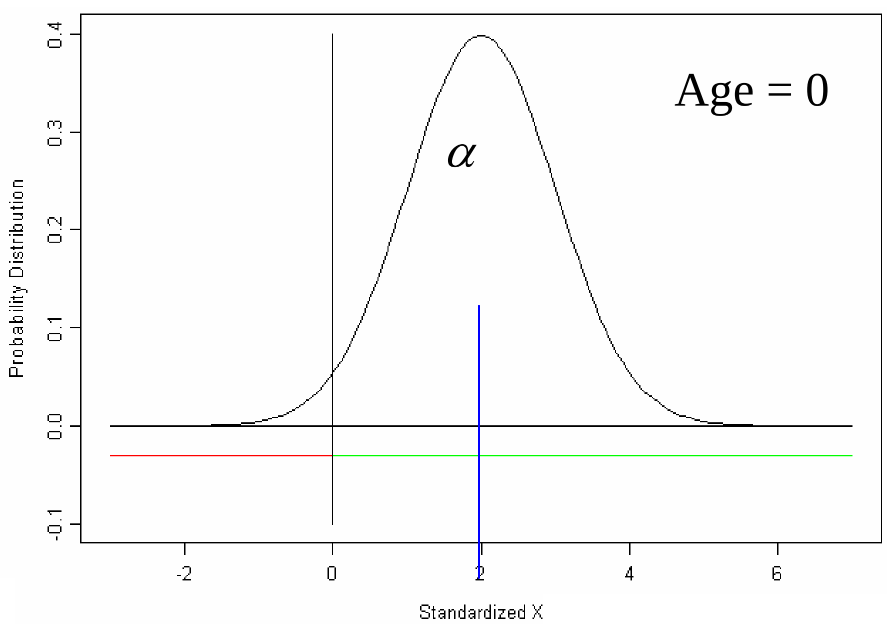
- To combine the data from these different data sources, we need an approach that allows flexibility:
  - Ability to incorporate expert knowledge of system
  - There is a considerably variability in how much data is observed for different pieces of the system
  - Not all components will have quality assurance data
  - The specification limits are thought to be approximations of when the part will fail, but do not necessarily match exactly with the flight data
  - Observed flight failure modes will not necessarily specify the failure of every component
  - There is frequently ambiguity about which component failed during flight testing

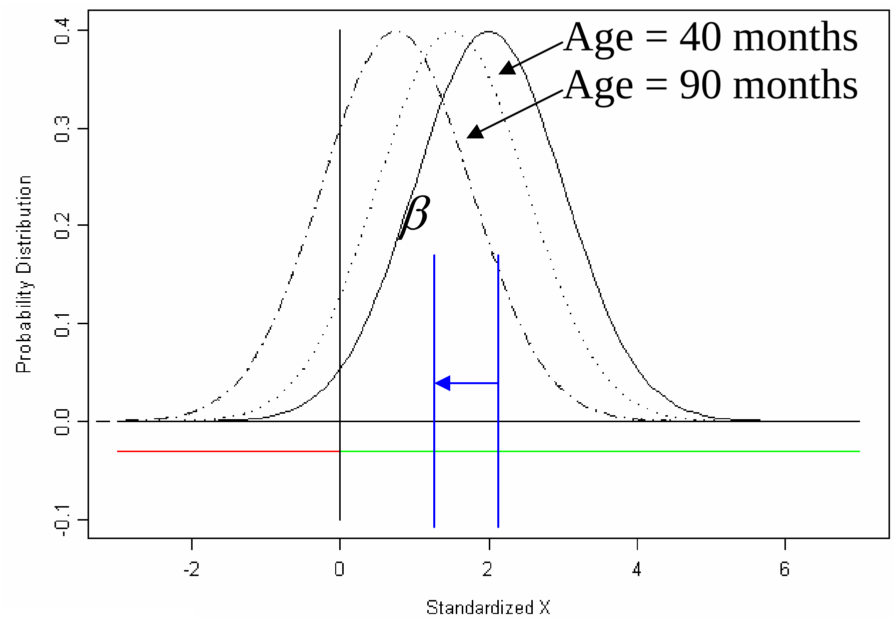
Including full-system data in the posterior distribution

- Define p_{1i} , p_{2i} , and p_{3i} to be the probability that components 1,2,3 work in the *i*th test
- These are functions of the age of the *i*th missile and of the unknown parameters, which we will define later
- For a very simple system with 2 components, we obtain terms like $(p_{1i}p_{2i})$ Both components worked
  $\{p_{1i}(1-p_{2i})\}$, Component 1 worked, but comp 2 failed
  and $\{1-p_{1i}p_{2i}\}$, At least one of Comp 1 or 2 failed
- For a more complex system, we might obtain
  $(p_{1i}p_{2i}\, p_{3i}p_{4i}\, p_{5i}p_{6i}\, p_{7i}p_{8i})$ All 8 components worked
  or $(p_{1i}p_{2i}\, p_{3i}p_{4i}\, (1-\, p_{5i}p_{6i}))$ At least one of C5 or C6 failed

Models for the QA measurements

- Denote the *i*th component QA measurement by C_i. It was taken from a missile with age A_i.
- Assume $C_i \sim N(\alpha_{Li} + \beta_{Li} A_i, \gamma_{Li}^2)$: linear regression
- α's have prior mean to match expected proportion of failures, β's should be negative
- Generates normal density terms in the posterior

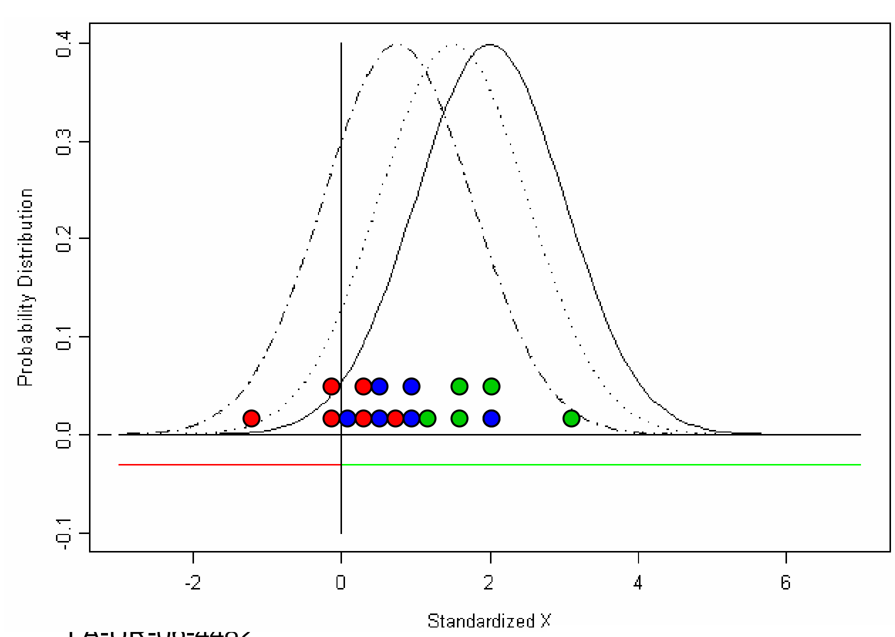




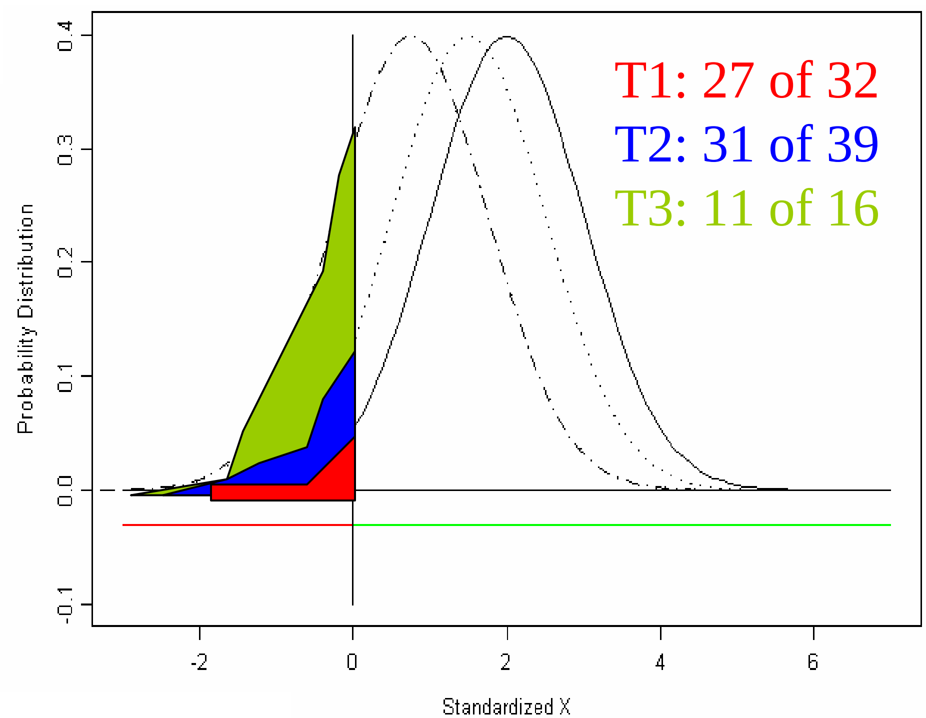
Comparing Sources of Data

- Both sources of data provide information about the shift of reliability over time

From QA data, we obtain the mean of the characteristic at each time



From the flight data, we obtain a proportion of success/failure at each time

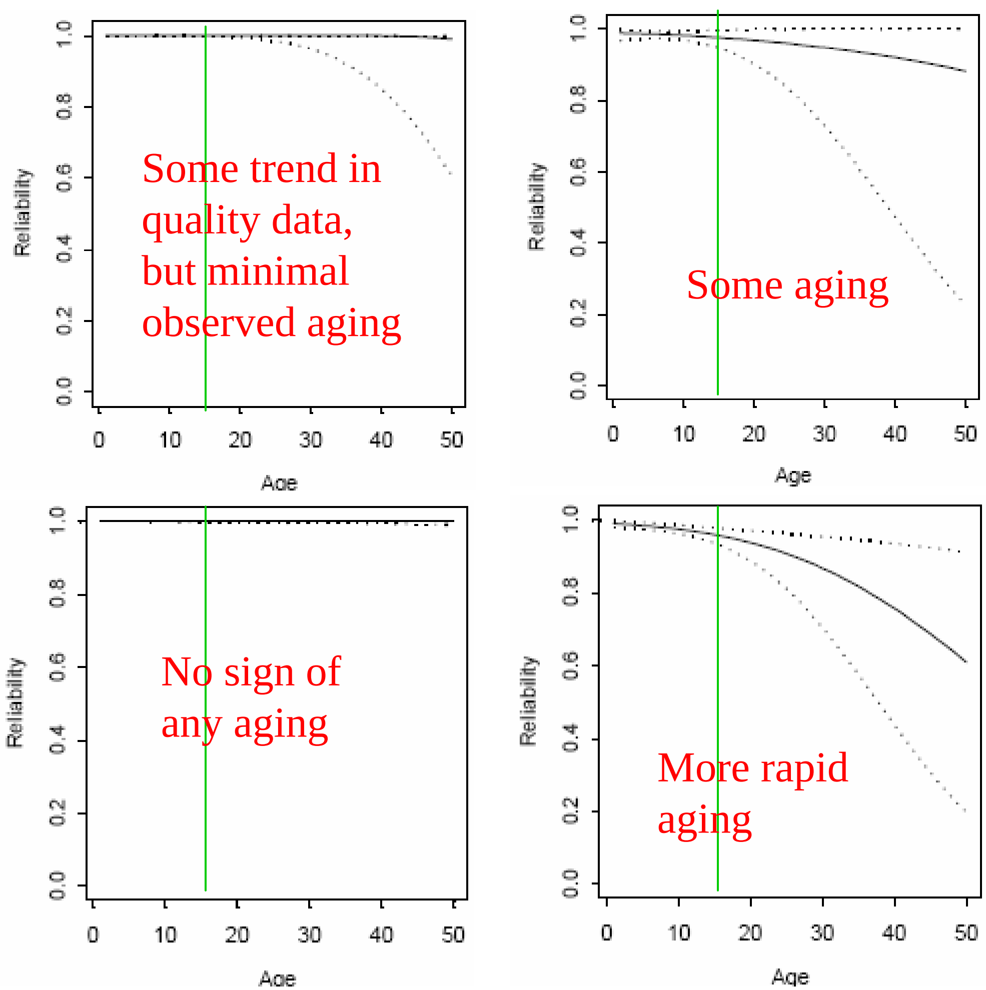


Model Analysis Outputs

- Component specific reliability estimates
  - For all observed times
  - For future times
- System level reliability estimates
  - For all observed times
  - For future times
- Information about how closely the current specification limits match what has been observed
  - This could be helpful for understanding the actual performance (i.e. what values of some of the quality assurance measures are actually associated with failures)

Component Reliability Estimates

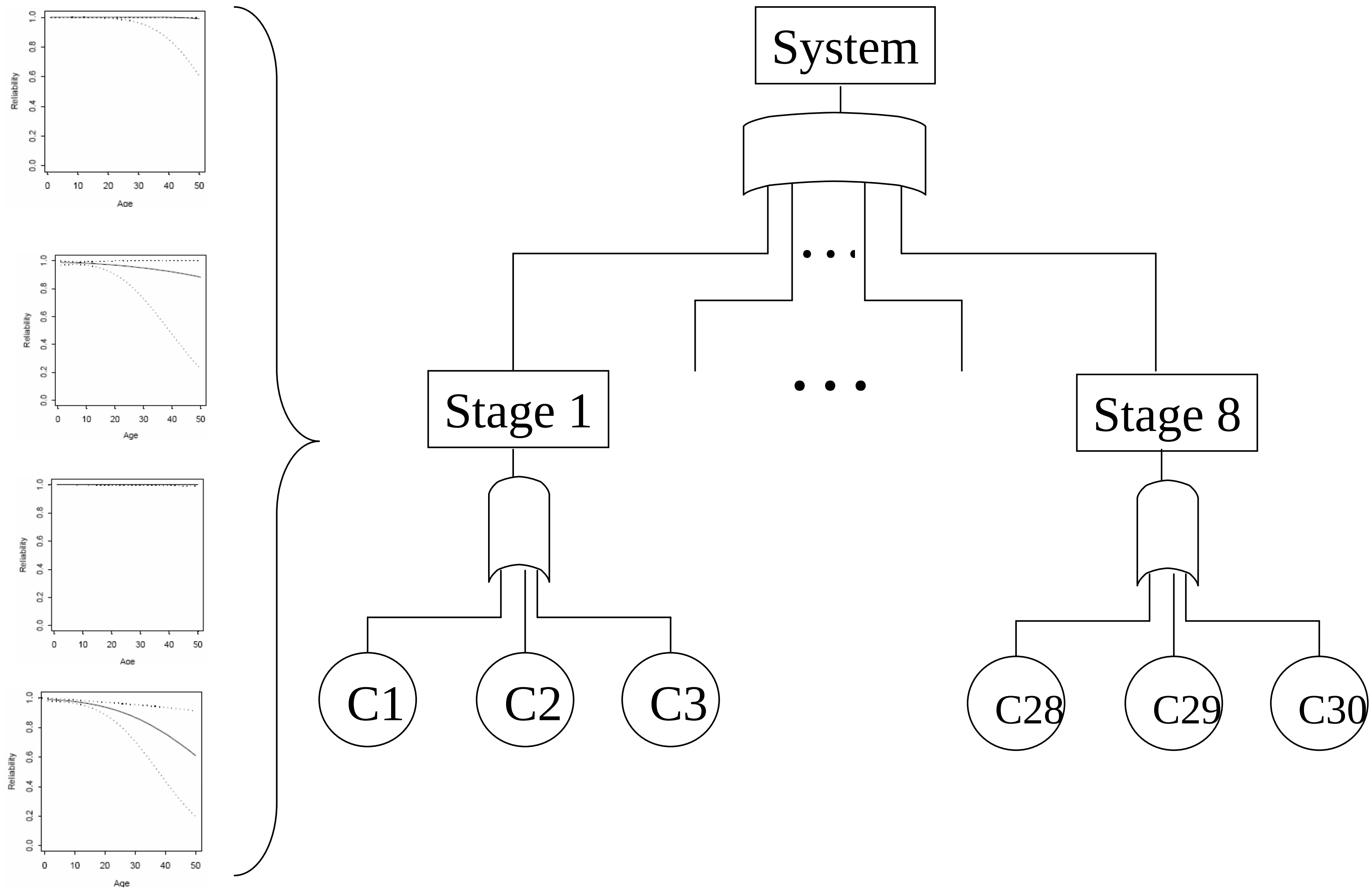
- For each component in the system, we can obtain estimates for its reliability at any age



Each component has its own summary with potentially different reliability and precision

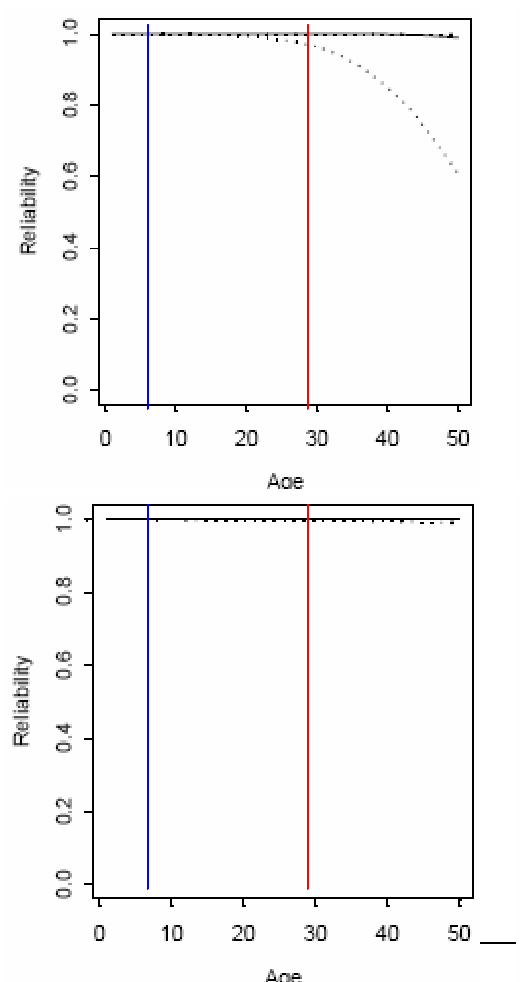
It is not uncommon to have many components showing little or no aging, while others are the main drivers of the system reliability

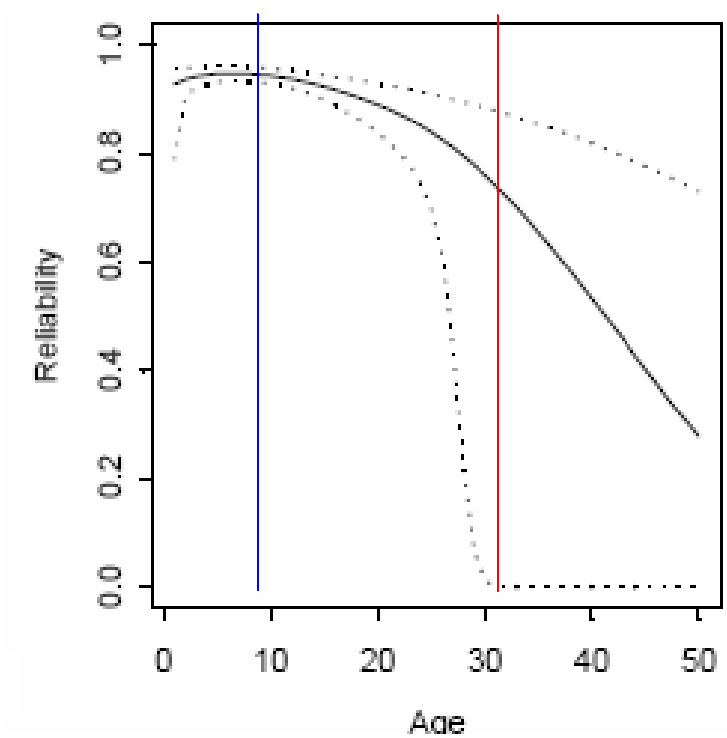
System Reliability Estimate



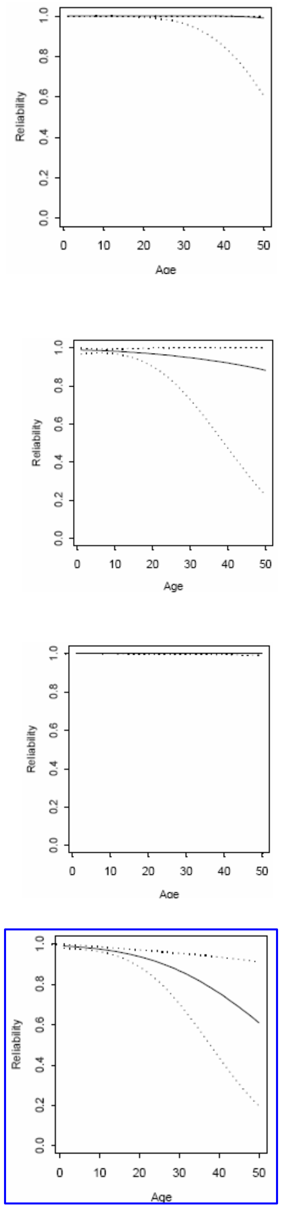
System Reliability at any age is the product of all of the component reliabilities in a serial system

P(system success) = function of component reliabilities



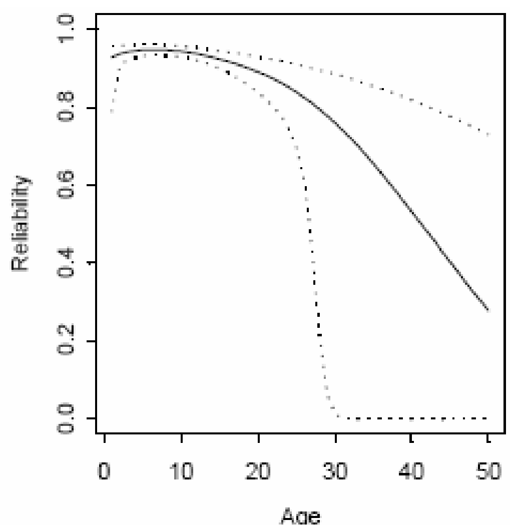


System Reliability for Variants

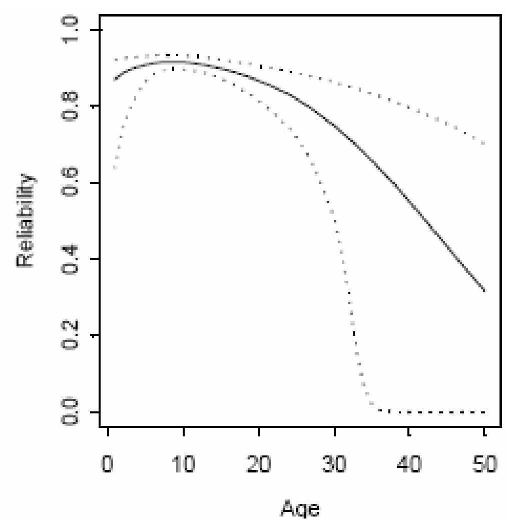


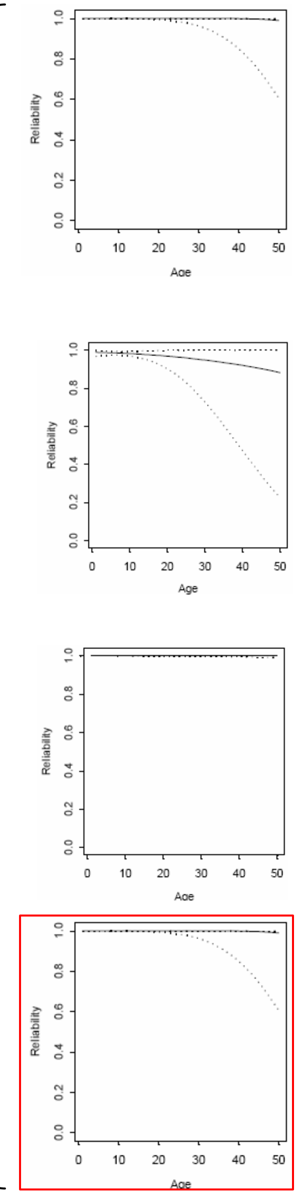
Recall for some systems there will be variants with many common parts, but some that are different. With this approach we can assemble a system estimate for any collection of components

Variant 1



Variant 2





Current Research

- There are many enhancements to the model which will make it increasingly flexible for variations in the data. We are currently working to include:
  - Incorporating additional system level covariates (e.g., Storage patterns, usage patterns)
  - More flexible types of quality assurance data (pass/fail, categorical, ordinal data)
  - Incorporating alternate data sources: maintenance, accelerated testing
  - Improving global summaries of stockpile reliability
  - Resource allocation strategies for collecting future data, based on current understanding of system and cost

SF = system flight
CF = component flight
CS = component spec (QA data)

Conclusions

- Modeling system reliability as a function of component reliability allows for additional sources of data to be included
- Extensive simulation study currently being conducted to help determine which system and data characteristics are most influential on accuracy and precision
  - Considered:
    - System complexity
    - Pattern of reliability over time
    - Number of variants
    - Amount and distribution of data over time
    - Amount and distribution of data between different data sources
    - Priors (diffuse, informative and correct, informative and incorrect)
    - Different data used: SF, SF CF, SF CS, SF CF CS

System characteristics

Data distribution

216 combos
x
10 reps

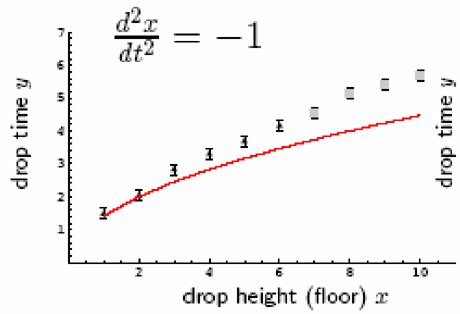
8 analyses

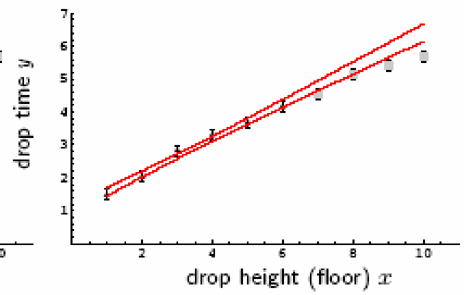
Conclusions (continued)

- There can be important advantages (for both accuracy and precision) to incorporating the component flight (CF) and component specification (CS) data
- Collecting component flight data is more beneficial for complex systems
- Priors need to be carefully chosen to reflect current understanding of component and system reliability (both diffuse and incorrect informative priors can cause problems)
- Cost considerations for the relative cost of collecting these data should also be considered when determining which analysis is best

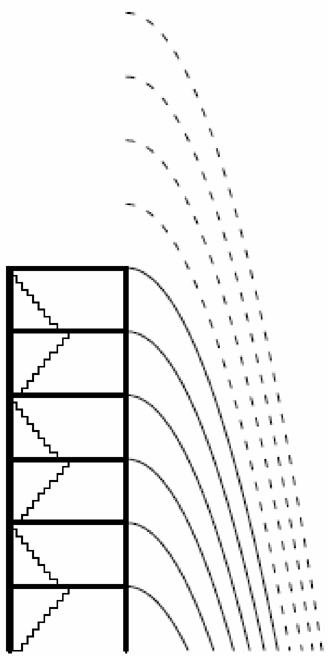
Inference Combining a Physics Model with Experimental Data

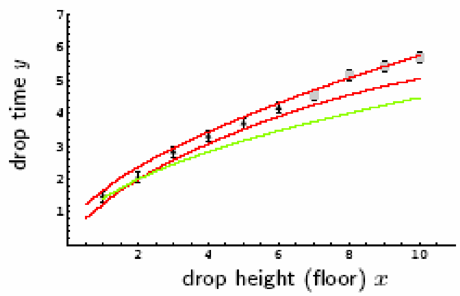
Data and simulation model:





Regression model:
$y(x) = \beta_0 + \beta_1 x + \epsilon$

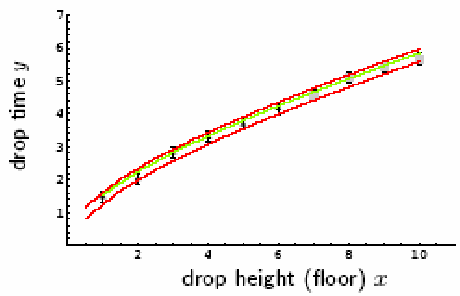




statistical/sim model:

$$y(x) = \eta(x) + \delta(x) + \epsilon$$

$$\eta(x) : \quad \frac{d^2x}{dt^2} = -1$$

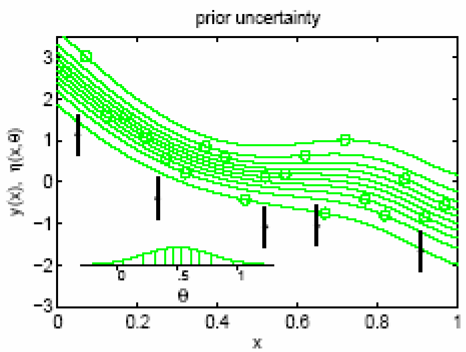


Improved physics model:
$\eta(x, \theta) : \quad \frac{d^2x}{dt^2} = -1 - \theta\frac{dx}{dt} + \epsilon$

statistical model:
$y(x) = \eta(x, \theta) + \delta(x) + \epsilon$

Statistical Framework

A statistical framework allows us to account for observational, experimental, and model errors



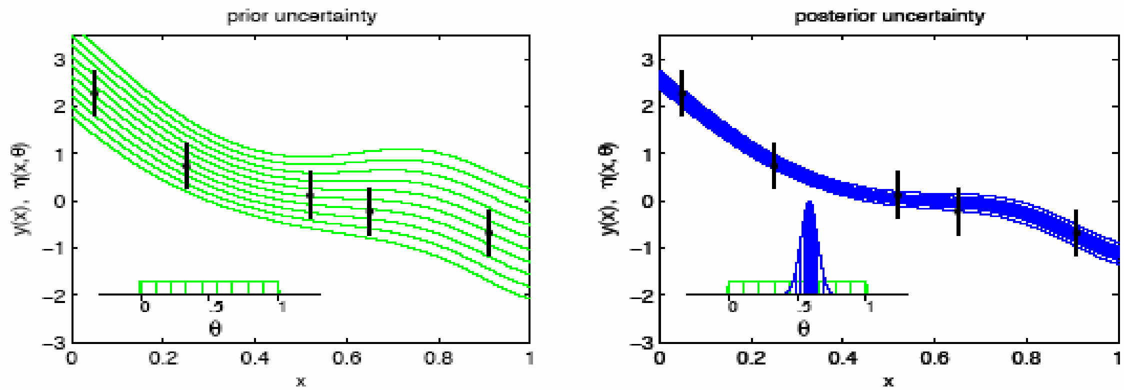
x	model or system inputs
θ	model calibration parameters
$\zeta(x)$	true physical system response given inputs x
$\eta(x,\theta)$	simulator response at x and θ.
$y(x)$	experimental observation of the physical system
$\delta(x)$	discrepancy between $\zeta(x)$ and $\eta(x,\theta)$ may be decomposed into numerical error and bias
$e(x)$	observation error of the experimental data

$$
\begin{aligned}
y(x) &= \zeta(x) + e(x) \\
y(x) &= \eta(x,\theta) + \delta(x) + e(x) \\
y(x) &= \eta(x,\theta) + \delta_n(x) + \delta_b(x) + e(x)
\end{aligned}
$$

Inference based on posterior

Uncertainty regarding θ, η, δ accounted for.

Statistical Formulation

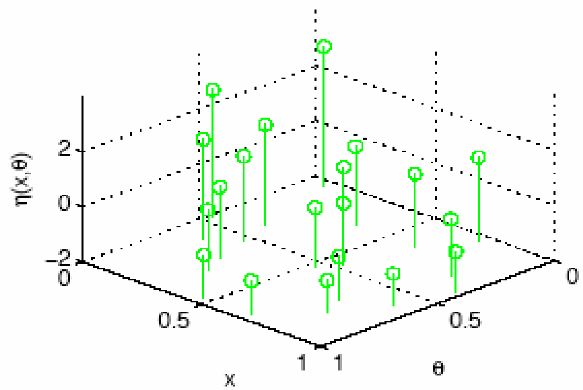


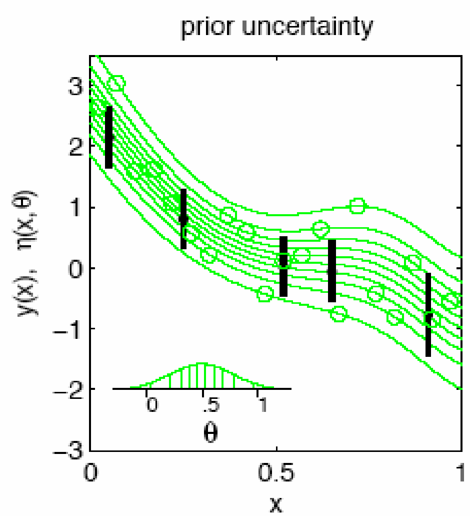
Observe data $y = (y_1, \ldots, y_n)^T$ at input conditions $x_1, \ldots, x_n$.

x	model or system inputs
θ	model calibration parameters
$\zeta(x)$	true physical system response given inputs x
$\eta(x, \theta)$	simulator response at x and θ.
$y(x)$	experimental observation of the physical system
$\delta(x)$	discrepancy between $\zeta(x)$ and $\eta(x, \theta)$
	may be decomposed into numerical error and bias
$e(x)$	observation error of the experimental data

$$
\begin{aligned}
y(x) &= \zeta(x) + e(x) \\
y(x) &= \eta(x, \theta) + \delta(x) + e(x) \\
y(x) &= \eta(x, \theta) + \delta_n(x) + \delta_b(x) + e(x)
\end{aligned}
$$

Experimental Design to Account for Limited Simulator Runs

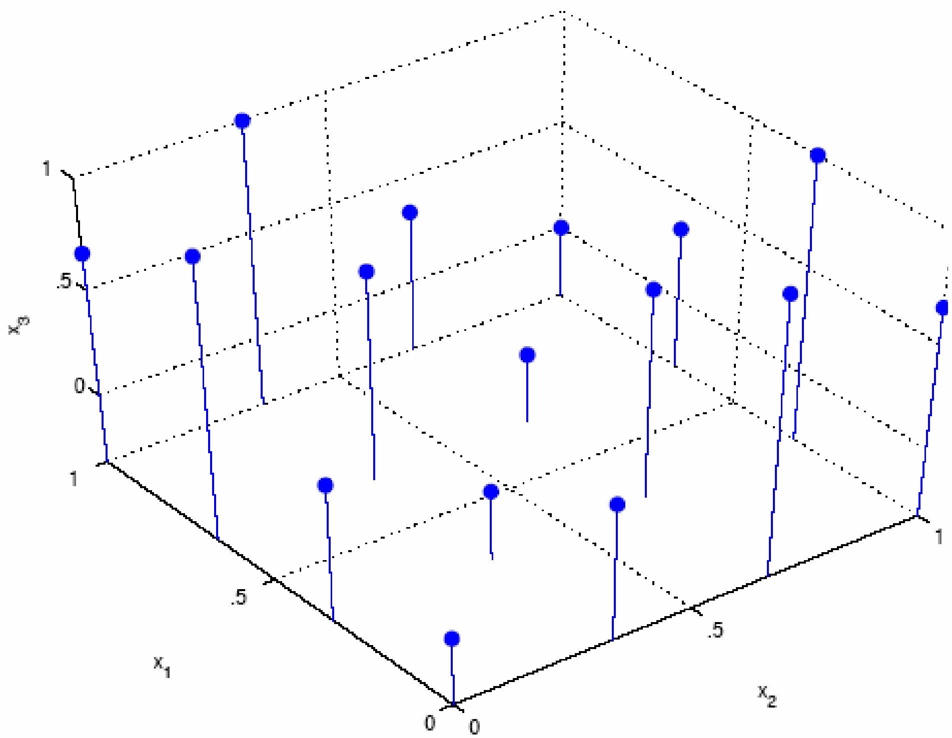


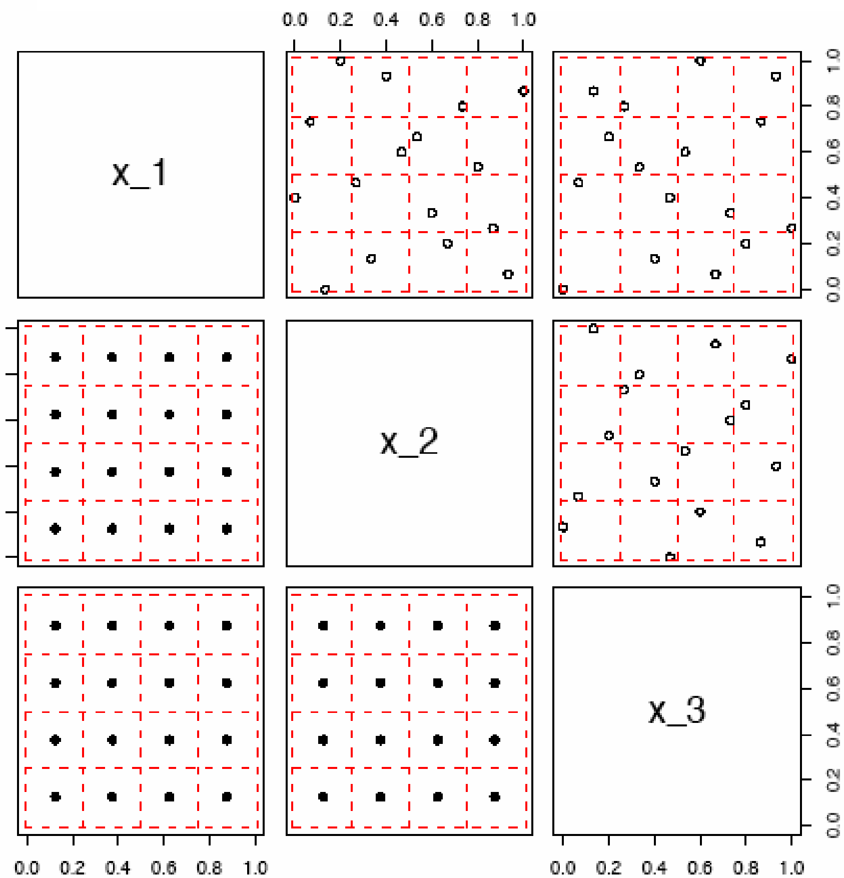


x	model or system inputs
θ	model calibration parameters
$\zeta(x)$	true physical system response given inputs x
$\eta(x,\theta)$	simulator response at x and θ.
	simulator run at limited input settings
	$\eta = (\eta(x_1^*, \theta_1^*), \ldots, \eta(x_m^*, \theta_m^*))^T$
	treat $\eta(\cdot,\cdot)$ as a random function
	use GP prior for $\eta(\cdot,\cdot)$
$y(x)$	experimental observation of the physical system
$e(x)$	observation error of the experimental data

$$
\begin{aligned}
y(x) &= \zeta(x) + e(x) \\
y(x) &= \eta(x,\theta) + e(x)
\end{aligned}
$$

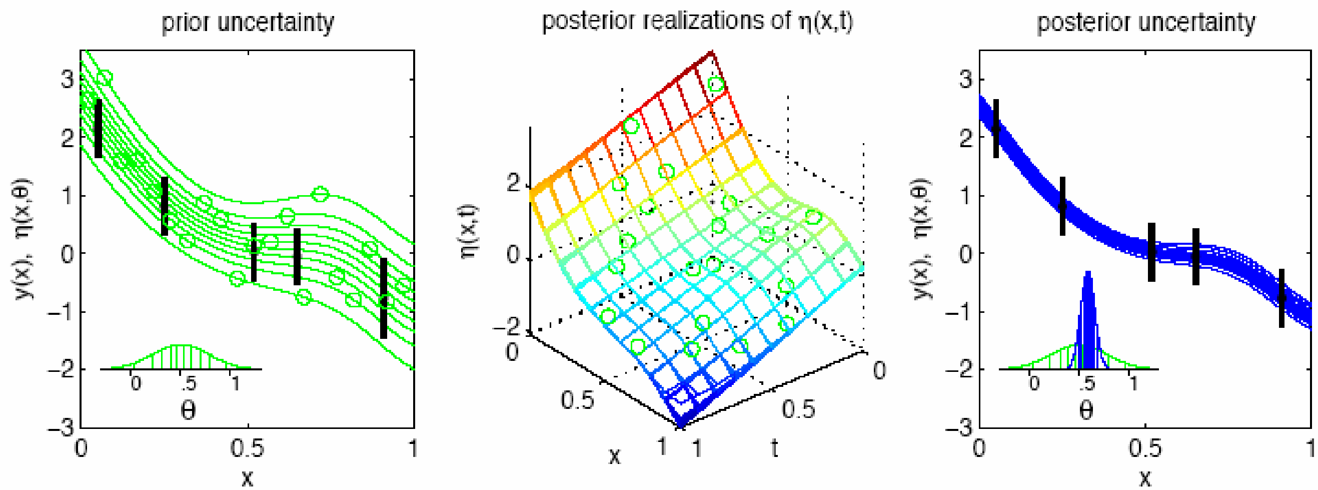
Experimental Design to Account for Limited Simulator Runs





Example: OA-based LH design

Modeling Simulator Response

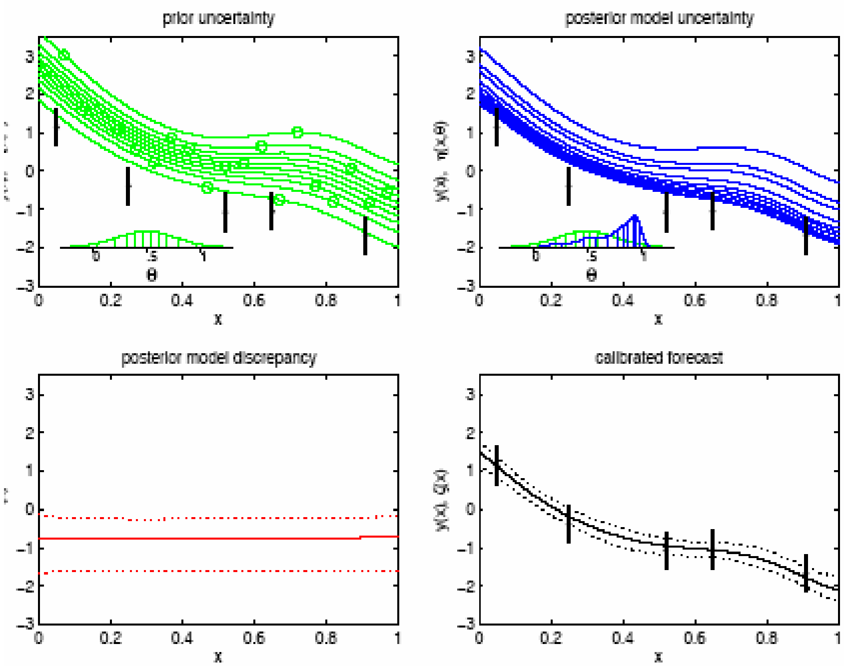


Again, standard Bayesian estimation gives:

$$\pi(\theta, \eta(\cdot,\cdot)|y(x)) \propto L(y(x)|\eta(x,\theta)) \times \pi(\theta) \times \pi(\eta(\cdot,\cdot))$$

- Posterior means and quantiles shown.
- Uncertainty in θ and $\eta(x,\theta)$ are incorporated into the forecast.
- Gaussian process models for $\eta(\cdot,\cdot)$.

Accounting for Model Discrepancy



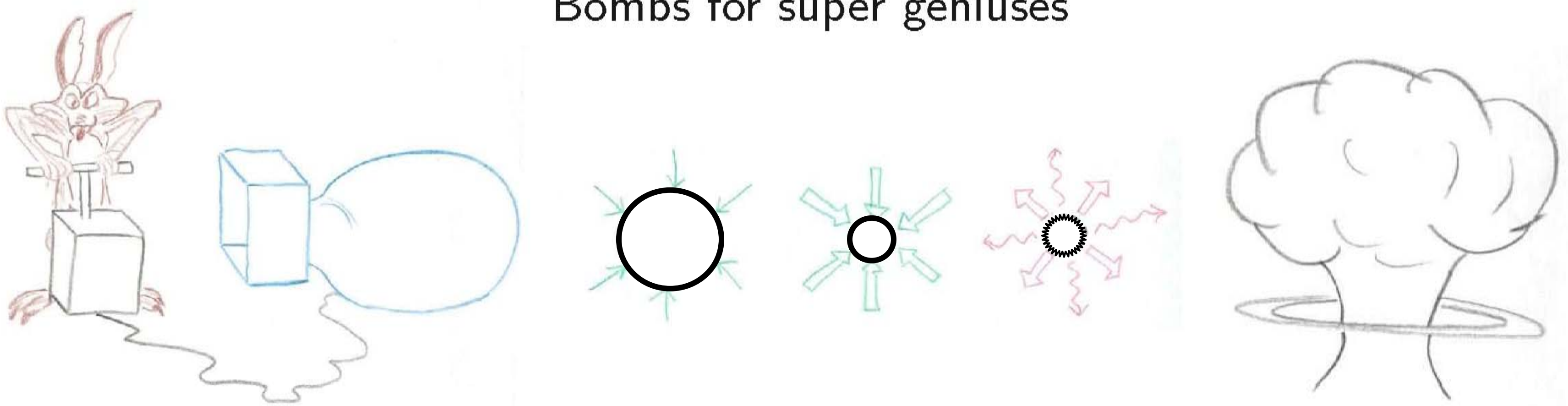
Again, standard Bayesian estimation gives:

$$\pi(\theta, \delta_n, \delta_b | y(x)) \propto L(y(x)|\eta(x,\theta), \delta(x)) \times \pi(\theta) \times \pi(\eta) \times \pi(\delta)$$

- Posterior means and 90% CI's shown.
- Posterior prediction for $\zeta(x)$ is obtained by computing the posterior distribution for $\eta(x,\theta) + \delta(x)$
- Uncertainty in θ, $\eta(x,t)$, and $\delta(x)$ are incorporated into the forecast.
- Gaussian process models for $\eta(x,t)$ and $\delta(x)$

Certification Issues at LANL

Bombs for super geniuses



time

detonation implosion nuclear yield

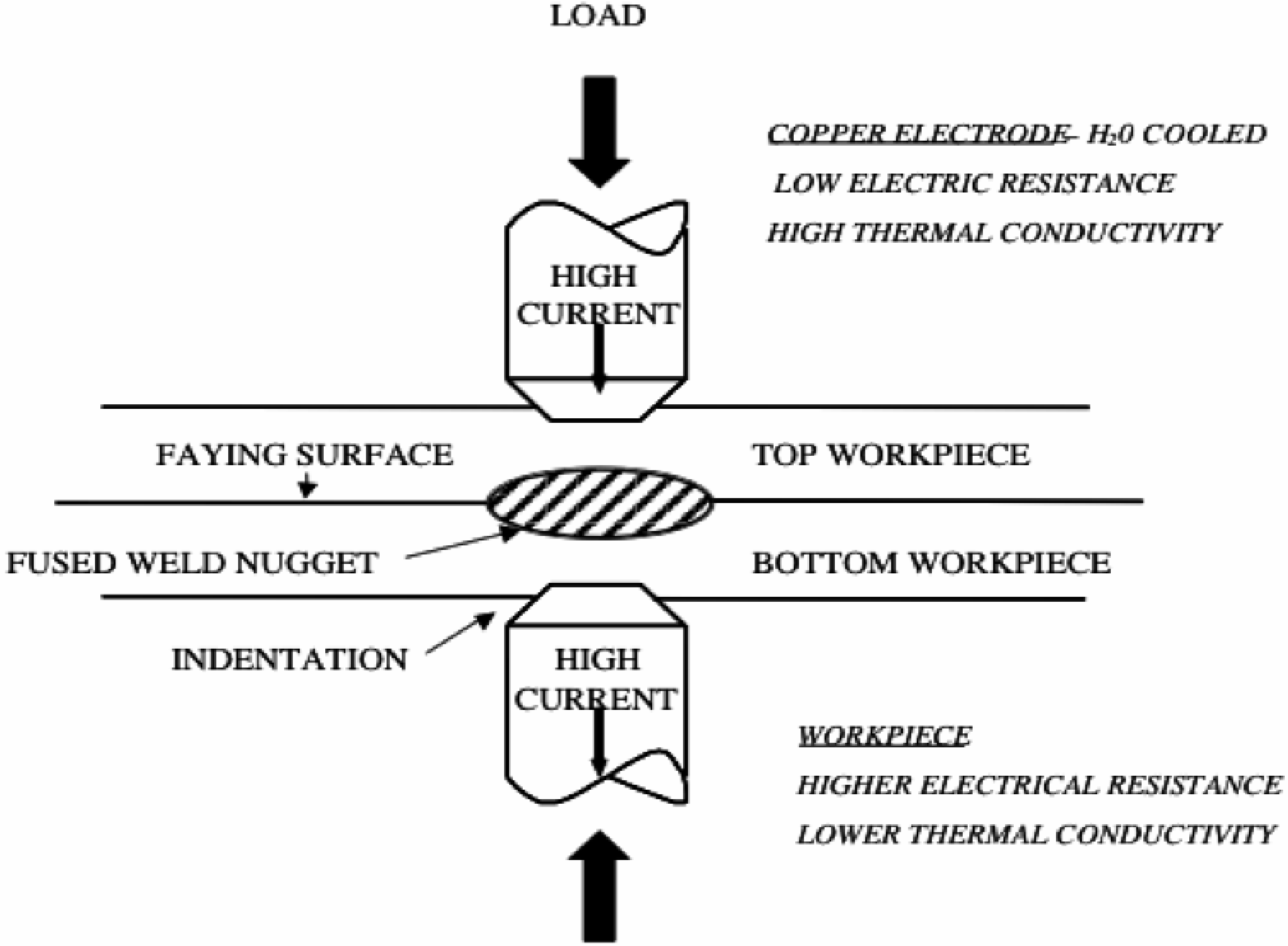
implosion experiments

sub-critcal experiments

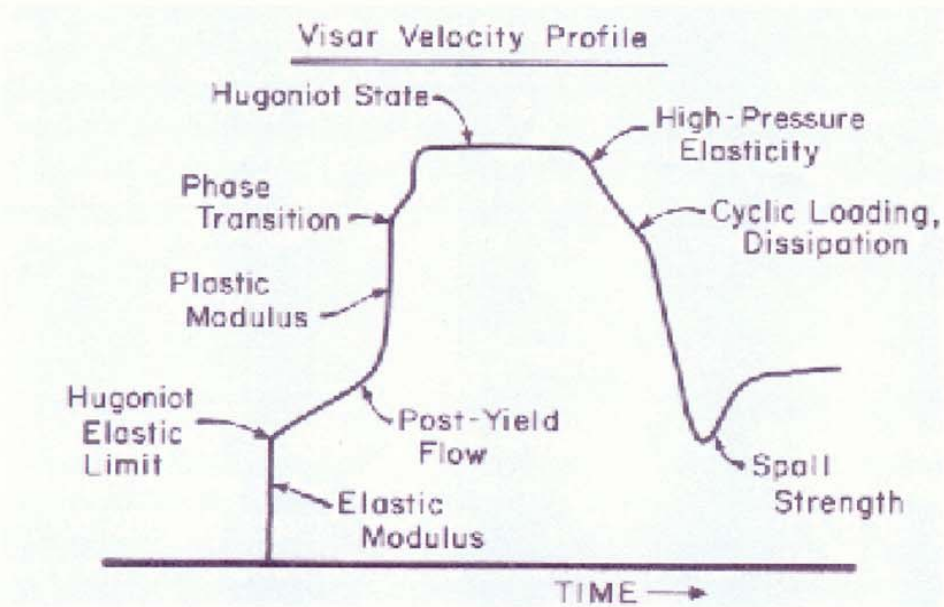
historical nuclear tests

- off-line experiments
  materials, equations of state (EOS), high explosive (HE)

Example: Spotwelding – Combining experimental data and simulations (with Marc Kennedy Univ Sheffield)



Calibration of Flyer Plate Calculations to Observational Data



- Velocity profile a function of material constitutive behavior
- Preston-Tonks-Wallace (PTW) model utilized in calculations to describe stress-strain relationship
- Calibrate free PTW parameters (7) to observational data

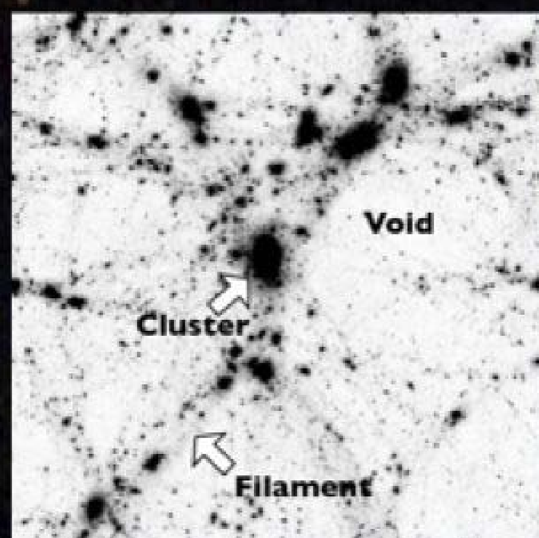
Cosmic Calibration

Dave Higdon, Katrin Heitmann, Charlie Nakhleh,
Salman Habib

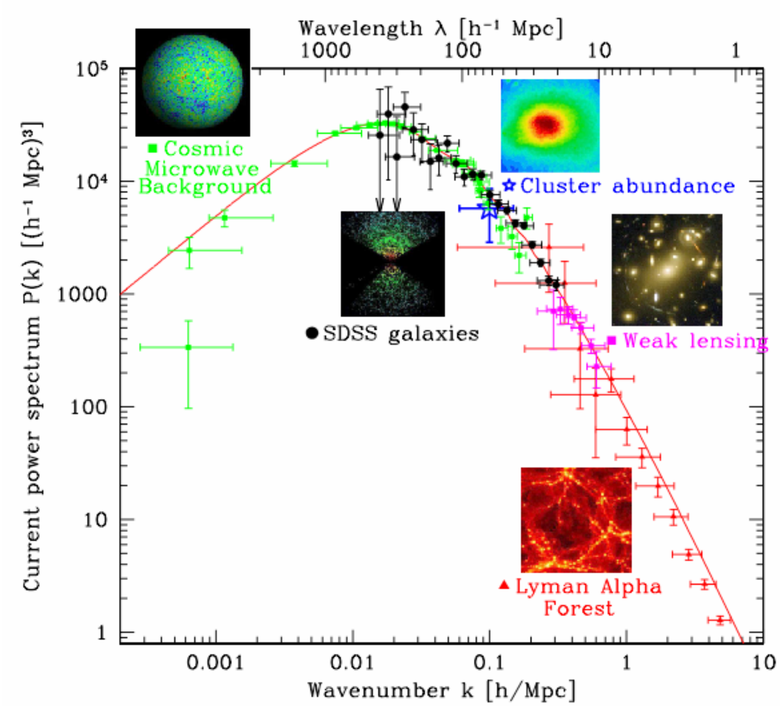
Los Alamos National Laboratory

SDSS

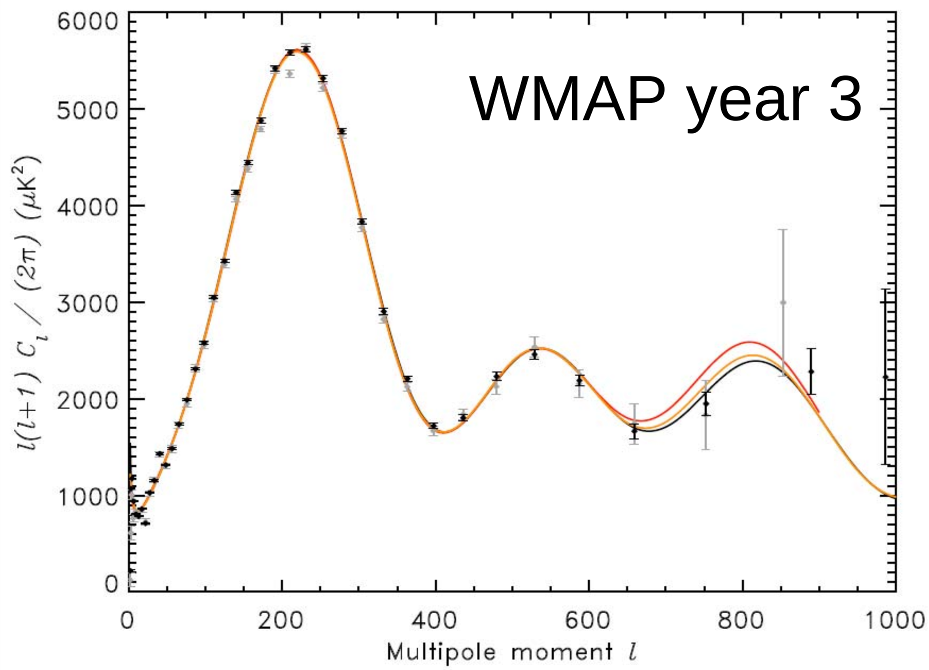
QSC



We want to estimate fundamental cosmological parameters using functional data from multiple data sources and simulations



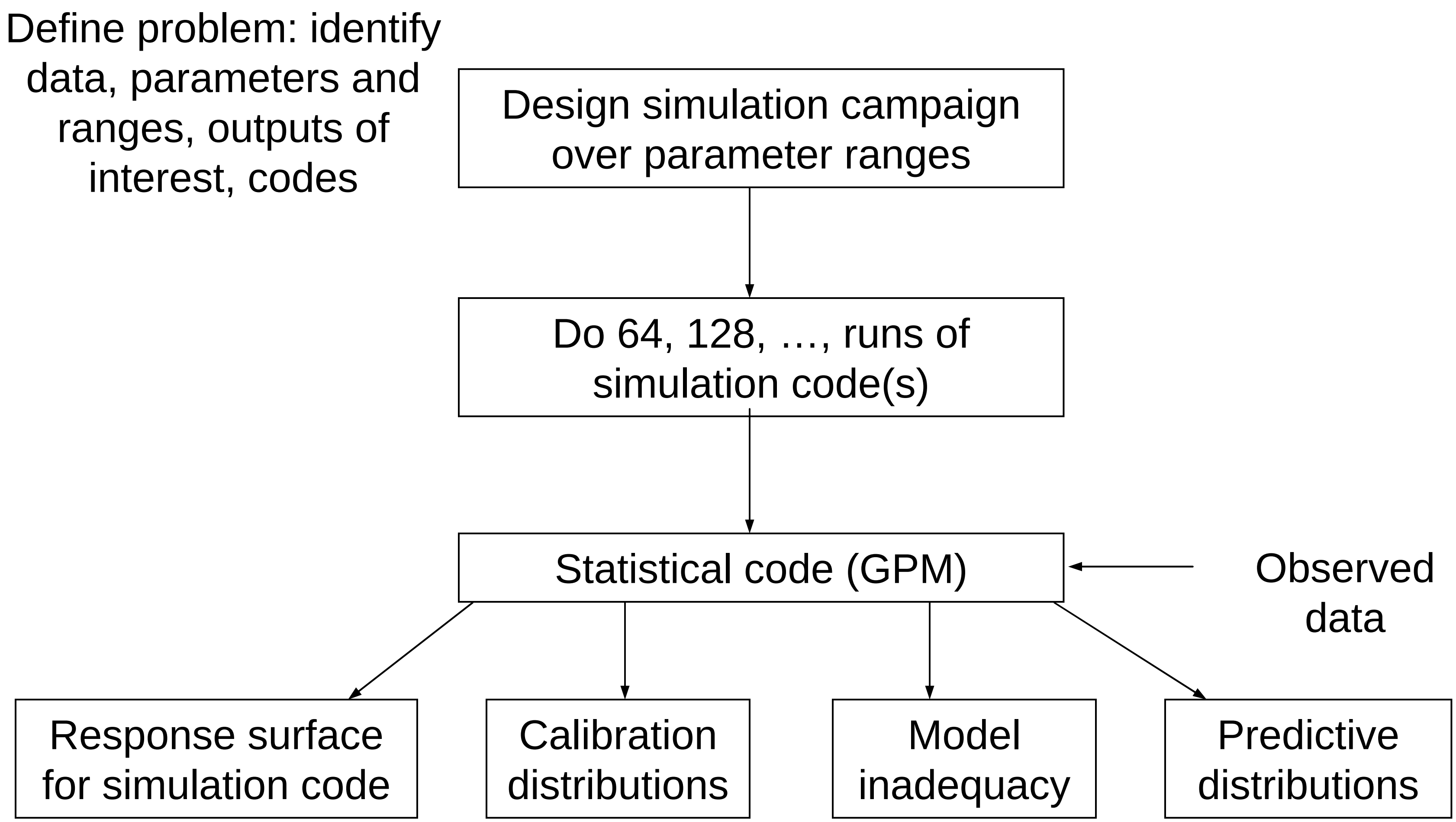
Linear spectrum only



Matter power spectrum
Linear theory
MC^2

CMB anisotropy
CMBfast

Our analyses use statistical methods to combine different simulation codes and observational data



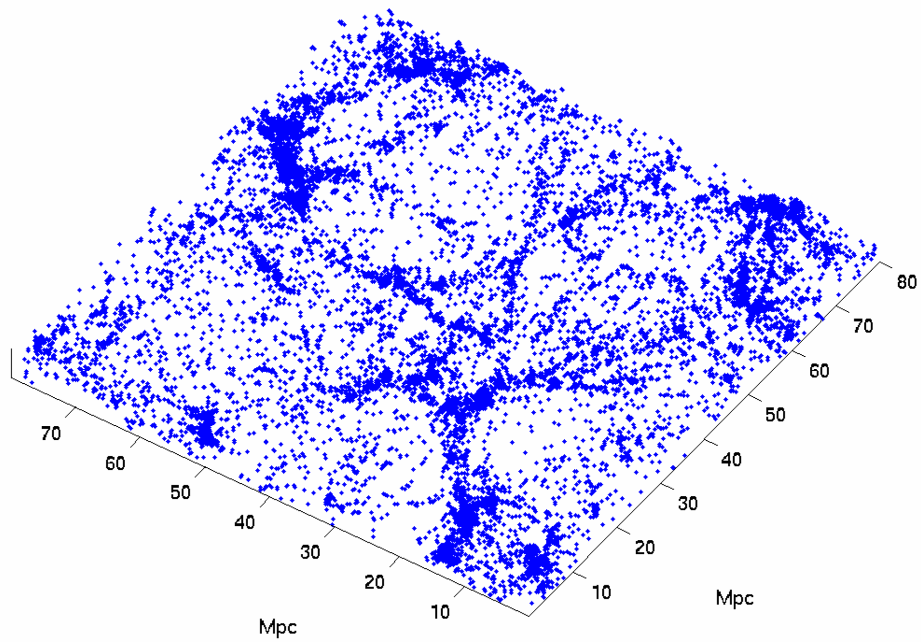
Data, parameter ranges, simulations

Log P

QuickTim
TIFF (LZW) d
are needed to se

Log k

A single simulation

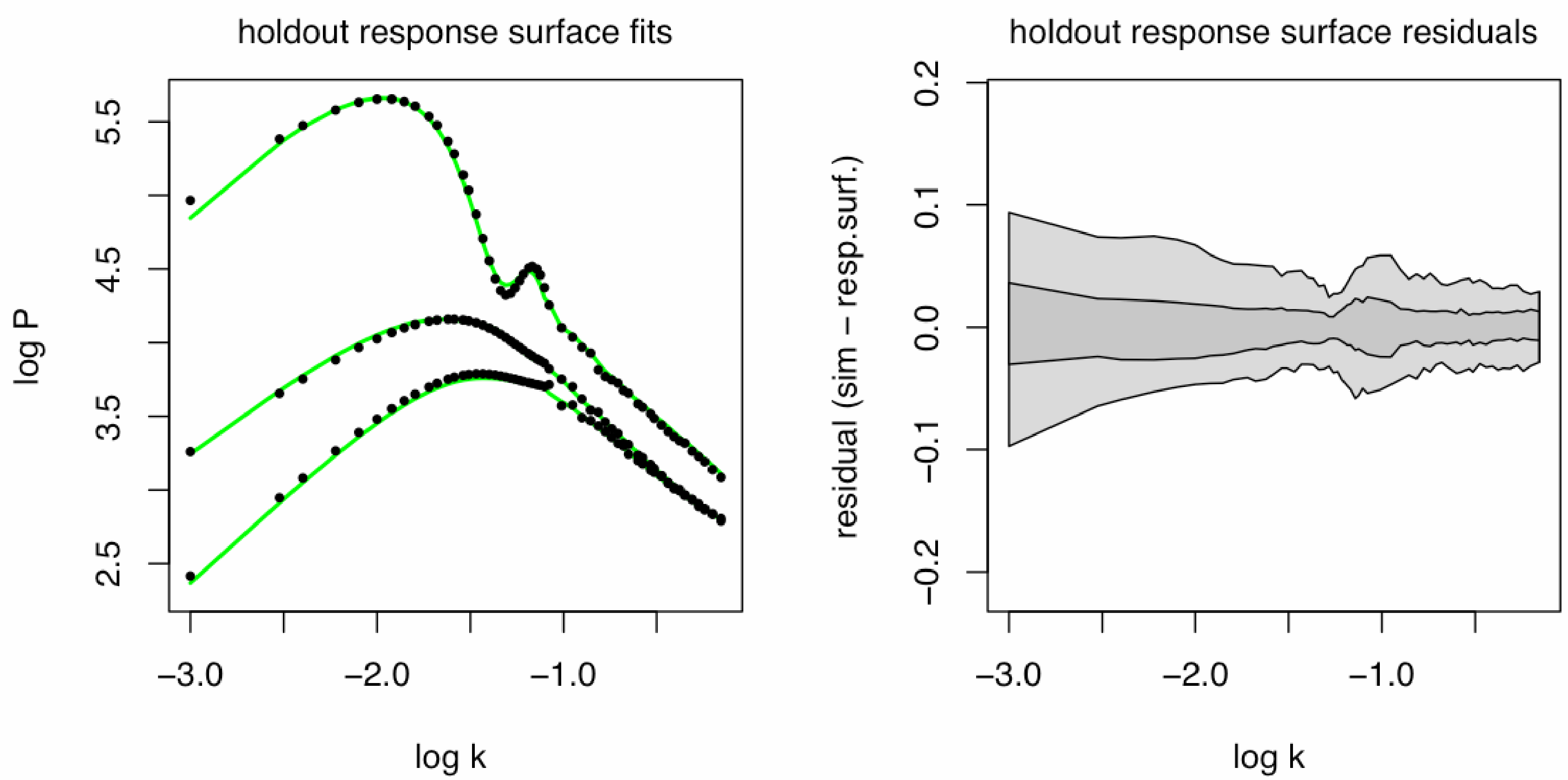


Synthetic data were generated from a “true” cosmology using both linear pertubation theory and the particle mesh code MC^2

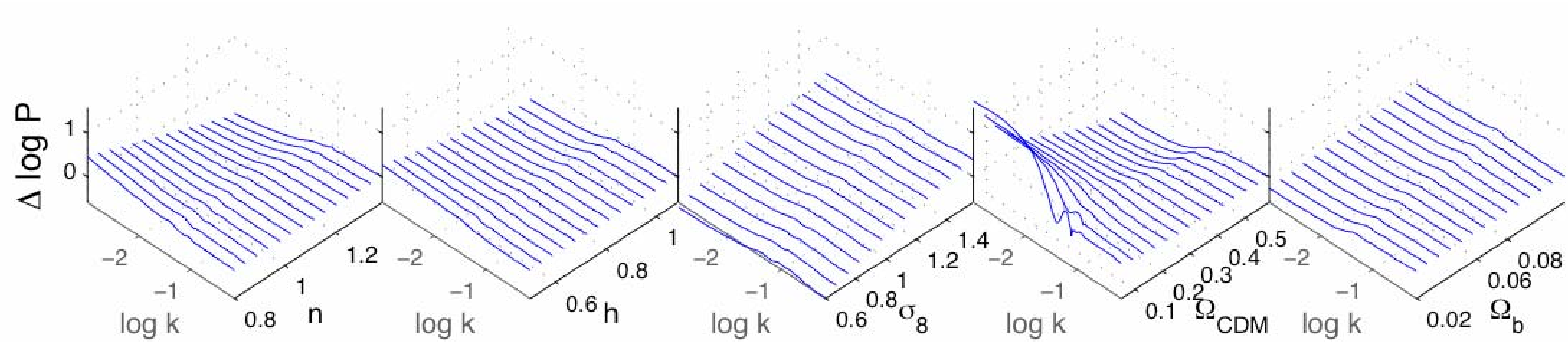
<u>Calibration parameter ranges</u>

Spectral index	0.8 to 1.4
Hubble parameter	0.5 to 1.1
Sigma 8	0.6 to 1.6
Omega CDM	0051 to 0.6
Omega baryon	0.02 to 0.12

Response surface accuracy



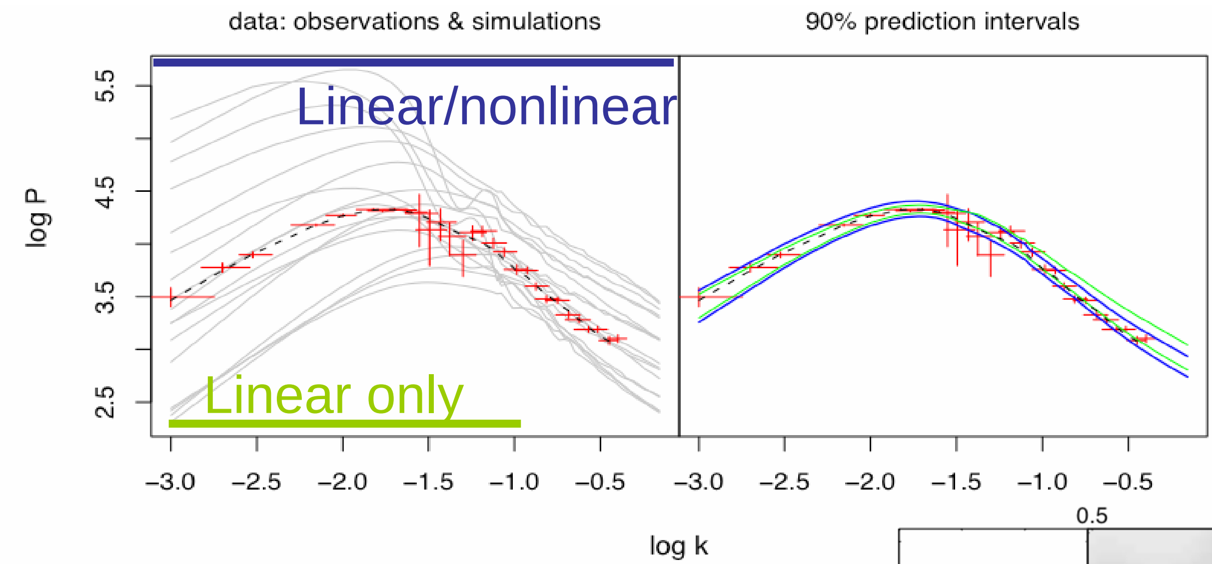
Simulator emulation and sensitivity



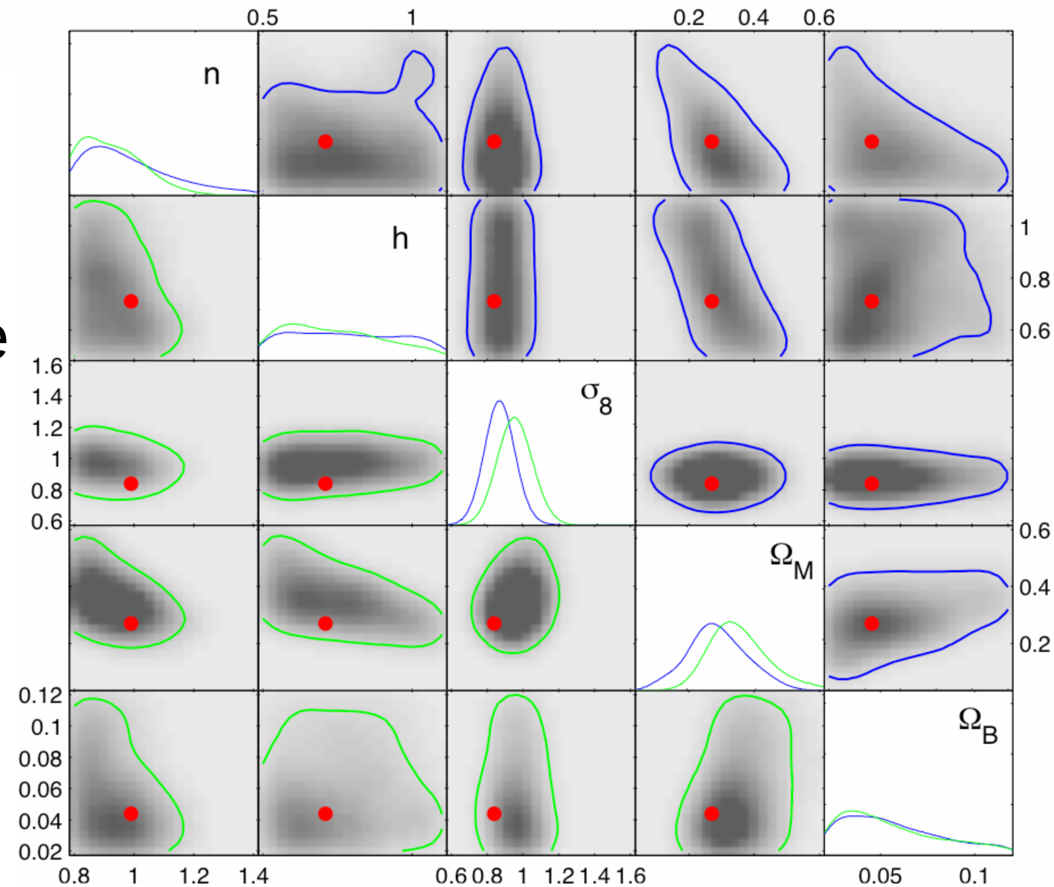
Changes in the emulator prediction as each parameter is varied while holding the others at their midpoint.

Note: σ_8 and Ω_{CDM} have the largest effect on log P

Only σ_8 has a substantial effect on nonlinear part of the mass power spectrum (logk < -1)



Posterior distribution



Two separate analyses:
- Using data which lie on the linear part of the spectrum
- Using data over the entire spectrum